

Intelligent Automation for the Trial Master File with AI-driven Applications

At IQVIA, we recognize that the proof of a new technology's useability occurs when we get past pilots and into production. More than five years ago we began a project to focus exclusively on emerging capabilities that could help us improve business automation across our suite of products. This resulted in harnessing our forty years of domain experience and **producing intelligent applications that can be inserted into capabilities** wherever they are required.

MAXIMIZE EFFICIENCIES INTERNALLY AND EXTERNALLY

Today, IQVIA has grown to have multiple intelligent applications in production that we use to maximize efficiencies internally, as well as offer externally to our customers through the IQVIA Clinical Data Analytics Suite. An example of an in-production capability enhanced by artificial intelligence and used by over 6,000 employees across the globe in multiple languages today is our Intelligent eTMF application.

Running at a capacity of approximately **14,000 documents a day across 1000+ clinical trials**, IQVIA utilizes artificial intelligence to automate and augment tasks that support optimization of document processing, embedding intelligence into the document pipeline for any eTMF system.

IQVIA'S INTELLIGENT eTMF APPLICATION:



Build of capabilities that can **provide an upfront, real-time content review** and a **quality check back** to the user to ensure only good quality content/scans are coming into IQVIA.



Digitization of content into standard structure allowing for scans, images, pdfs, docs to align and process for compare/version control/duplication management.



Extract any fields necessary through use of OCR (handwriting, embedded images) and Semantic Ontologies that standardize free text.



Run classifications of over 2000+ groups/sub-groups per DIA standards (and through our 10M+ historical domain content definitions).



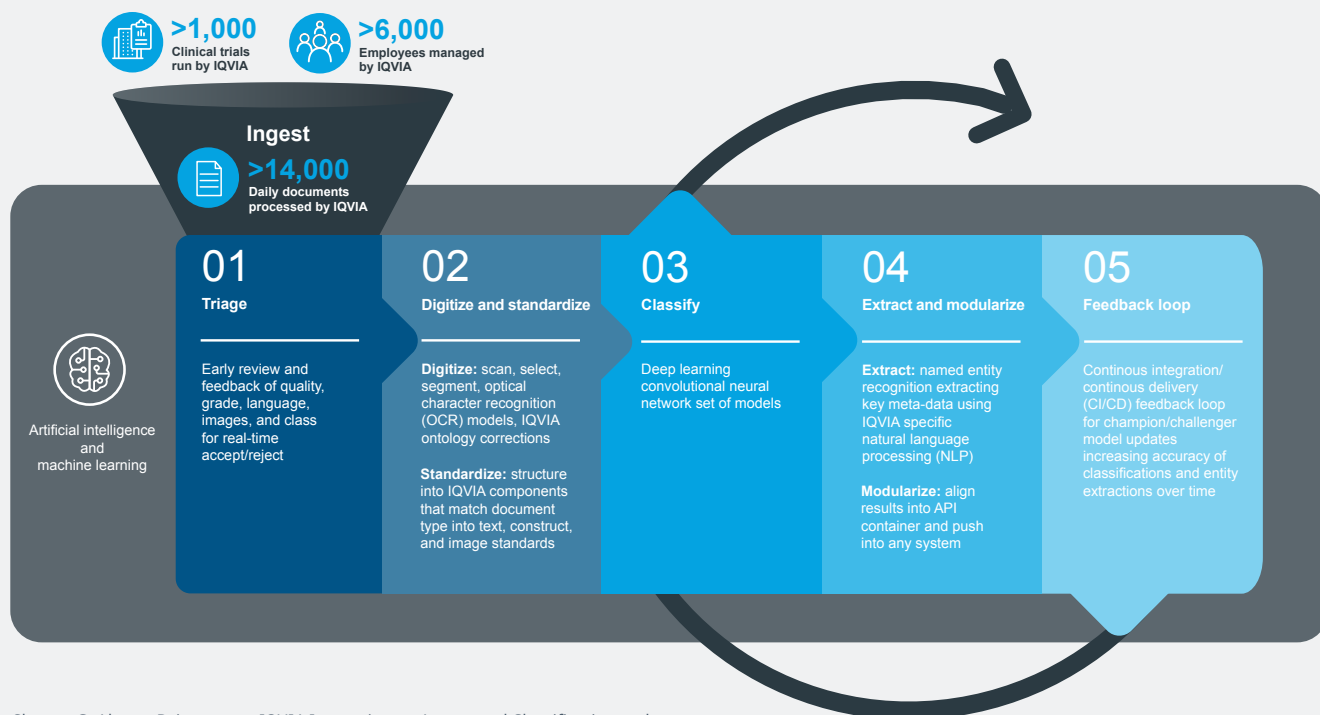
Feed the user the results for final confirmation and allow for their digital feedback to continue to evolve the accuracy of the classification and entity extractions.



Allow for integration (through our finalizer capability) **into any eTMF of choice**.

14,000 documents a day across 1000+ clinical trials, IQVIA utilizes artificial intelligence to automate and augment tasks.

Intelligent eTMF™



Shorter G, Ahrens B, inventors; IQVIA Inc. assignee. Automated Classification and Interpretation of Life Sciences Documents. US patent 10,839,205. November 17, 2020.

IQVIA values the user experience, both internally and with our clients, as a key part of the automation and workflow enhancement within our technology solutions. We view AI-driven intelligent applications working alongside humans to augment their experience - but without assuming a black box input/output mentality.

Our intelligent applications work to the latest regulatory standards to allow for a transparent experience for clients to understand how they were built, what software was used, what data helped to train the technology, how accurate the AI-driven application is and identify confidence intervals at the field level across all content.

Globally, in any country and in any language, from standard content to scans and images coming in from the sites – content must be ingested and reported. IQVIA recognizes the reality of how messy the content aggregation can be and offers best practices for a successful outcome.

1. TRIAGE

Triage is a necessary key in the early identification of content engagement. Certainly, standard format documents have a structure and hence are easy to extract into standard fields for recognition by a typical eTMF platform. However, IQVIA is capable of handling situations where users are remote, at sites and using whichever tools are at their disposal to bring the content into the system.

As such, having an early **Quality Check intelligent application identifying the type of content, recognizing a scan, running checks over that scan for completeness** (pagination, layout, viewability, language etc.) and either returning a 'proceed' to the next intelligent application or a text/e-mail back to the user regarding their document (whilst they are still in the field in real-time) can be invaluable to the progress needed to successfully manage a quality eTMF.

2. DIGITIZE AND STANDARDIZE

Subsequently, once a piece of content is recognized a trigger is sent to the right intelligent application for the job. As such, there are several that can be used, specific to the Triage results. With scanned content, the intelligent application would run through our AI trained OCR “AI Bots” that recognize specific domain content (including languages for accurate digitization) and can therefore digitize into standard structure for downstream extraction of key fields. This helps further with the **automation and accuracy of the extraction of relevant fields that an eTMF system requires.**

3. CLASSIFY

An intelligent application that can be built using a machine learning model is the classification model. Despite its usefulness, there are several challenges to overcome with classification. Firstly, the intelligent application must classify >10M records of historical content showing how >2000 content types have previously been denoted. Secondly, augmentation of that classification is necessary with as many unique field identifiers marked to be utilized within the machine learning model for the highest accuracy. This allows for artificial intelligence techniques to not rely on a sole field but make use of a multitude of fields from the digitization to help ‘recognize’ the content.

Most importantly to the success of AI adoption and usage is the level of accuracy amongst data flows. For each run, IQVIA provides a confidence score to the user to help them decide to either focus on a low score or push through content with high scores. These scoring models were developed because the machine learning model created is not just one model but a collated management of models. Typically, we have several levels of classification to account for from distinct groups of content down to more niche and smaller sample sizes of very specific classifications. IQVIA has focused on the broadest set of classifications to meet DIA standards available, and combined with our domain history, has resulted in a strong, mature state today.



4. EXTRACT AND MODULARIZE

With extraction and modularization of digitalized content, it is now possible for relevant fields required by the eTMF to be auto-inserted into an Intelligent eTMF-recognized format for consumption of the content alongside this meta-data. This allows the user to bulk load, run the intelligent application, collect all relevant fields in one location, update/edit where they feel they need to and then push the results straight into the eTMF, a marked improvement in process automation.

5. FEEDBACK LOOP

It is important to have a partner that understands that technology works best when it is easy to use. The IQVIA user interface allows for bulk load of content, and the easy-to-run Intelligent eTMF application with the ‘push a button’ feature enables the return of results and the ability to edit prior to pushing to the eTMF.

As such, any human update is automatically recorded and saved back to the intelligent applications harnessing data. This allows IQVIA to monitor across **>3M content collections in an Intelligent eTMF per year**, compare the human to AI results and run champion-challenger machine learning model version updates per good GXP practice. This unique feature provides the ability to improve accuracy over time as the intelligent applications continue to learn.

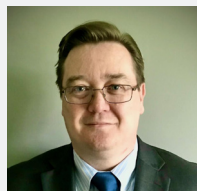
With these parameters firmly in place, IQVIA has shown through its own discipline that it can empower life science organizations to harmonize complex and disconnected data and **use AI/ML to draw smarter insights that improve clinical research outcomes for patients, sites, and sponsors.**

With a cloud-based, modular platform that anchors the clinical trial lifecycle by ingesting and standardizing previously disconnected operational data for inquiry, IQVIA has harnessed intelligent data aggregation, review, and evaluation.

The vendor-agnostic Clinical Data Analytics Suite platform offers customers a single, scalable repository for all clinical stakeholders to explore their operational and clinical data, assisted by AI. Customers derive actionable insights using analytics tools including for biostatistics and data science, rich visualizations, reporting with pre-built KRIs and KPIs and through the use of API micro-service intelligent applications.

At IQVIA, we understand both the practical application for the user and how maintaining a strong control of the data flow is important. Successful organizations must plan for business continuity whilst having the ability to run intelligent applications that support higher and more accurate quality efficiency gains going forward, including an intelligent eTMF document pipeline.

To learn more about our Intelligent eTMF™ application capabilities and other Intelligent Applications in our Clinical Data Analytics Suite, contact us for a demonstration at **OrchestratedYourTrials@iqvia.com**



GARY SHORTER
Head, A.I. and Data Science,
IQVIA Technologies

In IQVIA Technologies, Gary pursues the use of emerging technology to provide new and more efficient capabilities within the running of Clinical Trials. This included the development of new design software along with more recent advancements with AI/ML capabilities where his team developed and manages micro-products and services that can be plugged in and used by any SaaS solution to improve efficiency and decision-making capabilities.